SUMMARY

The main subject of this thesis is the biogenesis of microbodies in yeast. The methylotrophic yeast *Hansenula polymorpha* offers an attractive model system to study microbody biogenesis. The proliferation and enzyme content of microbodies is entirely dependent on the growth conditions and, hence, can be easily manipulated. As a first step towards the identification of topogenic signal(s) for protein import into the microbodies of this yeast, we chose to compare the amino acid sequences of a number of microbody enzymes derived from the nucleotide sequences of their genes. Putative topogenic signal(s) emerging from this comparison will then be tested in subsequent studies.

In Chapter 1 a general introduction dealing with the metabolic significance and biogenesis of microbodies in yeasts is presented. In addition, this chapter gives a brief overview of protein targeting signals and the translocation machinery, with emphasis on targeting signals that mediate protein transport into microbodies. Chapter 2 describes the isolation and characterization of the gene coding for amine oxidase, key enzyme in the metabolism of alkylated amines. The amine oxidase (AMO) gene contains an open reading frame of 692 amino acids, with a relative molecular mass of 77,435. Amine oxidase contains the tripeptide sequence Ser-Arg-Leu at 9 residues from the carboxy terminus. This sequence conforms to the conserved tripeptide motif Ser/Ala/Cys-Lys/Arg/His-Leu that, at least in one case, is required for protein targeting to the microbodies and has been found in several microbody proteins at or near the carboxy terminus. Therefore, we consider this sequence to be a likely topogenic signal for import of amine oxidase into the microbody. Comparative analysis of the primary structures of three microbody proteins of *H. polymorpha*, viz. amine oxidase, alcohol oxidase and dihydroxyacetone synthase, did not reveal any other common sequence that could be implicated as a possible targeting signal. We found two genes in the vicinity of the AMO gene. An analysis of these genes is presented in an appendix to Chapter 2. Chapter 3 describes the purification of malate synthase, one of the key enzymes of the glyoxylate cycle. Native malate synthase consists of four identical subunits with a relative molecular mass of approximately 62,000. We compared the enzyme with malate synthases from other fungal sources. Biochemical and immunocytochemical studies show that malate synthase is located in the matrix of the microbodies. The cloning and sequencing of the gene encoding malate synthase is described in Chapter 4. The malate synthase (MAS) gene contains an open reading frame of 555 amino acids, corresponding to a calculated molecular mass of 63,254 for the encoded protein. Comparison of the amino acid sequence with malate synthase sequences of *Escherichia coli*,

Brassica napus L. and *Cucumis sativus* L. shows a high degree of similarity between these proteins. However, the malate synthases destined for translocation across the microbody membrane differ from the *E. coli* enzyme by a short carboxy-terminal extension. In the plant malate synthases, this extension probably mediates the routing to the microbodies, since it contains the identified microbody targeting signal, Ser-Arg/Lys-Leu, at the carboxy terminus. Interestingly, malate synthase of *H. polymorpha* contains the same amino acids in a different arrangement, Ser-Leu-Lys, at the carboxy terminus. This result suggests that degenerate forms of the conserved tripeptide motif, when placed in the right context, are recognized by a microbody receptor. Future studies, which include gene transfer experiments combined with gene fusions and site-directed mutagenesis, should provide experimental evidence that the putative topogenic signals in amine oxidase and malate synthase are required for import of these proteins into the microbodies.

During growth of *H. polymorpha* on glucose/diethylamine, the acetaldehyde generated in diethylamine oxidation is further metabolized by aldehyde dehydrogenase. Yeast contains at least two isoenzymes of aldehyde dehydrogenase: an NADP-linked enzyme found in the cytosol, and an NAD-linked enzyme located in the mitochondrion. The isolation and sequencing of the gene encoding an aldehyde dehydrogenase (ALD) is described in Chapter 5. The nucleotide sequence of the ALD gene contains an open reading frame of 518 amino acids, corresponding to a calculated molecular mass of 56,142 for the encoded protein. Comparison of the derived amino acid sequence of the *H. polymorpha* ALD gene with aldehyde dehydrogenase sequences from other species clearly establishes the homology of all proteins. From comparison of the amino-terminal amino acid sequence of *H. polymorpha* ALD to the cleavable amino-terminal signal sequences of mitochondrial enzymes and particularly to the signal sequences of human and rat mitochondrial aldehyde dehydrogenase, we infer that we have cloned the gene encoding the mitochondrial isoenzyme. Definite proof can be obtained by amino acid sequence analysis of the purified enzyme.